

WHAT IS CLAIMED IS:

1. A device for interpreting XML documents, comprising:
 - an acquiring unit that acquires an XML document, wherein the XML document includes a plurality of elements that form a
 - 5 predetermined tree structure;
 - an arranging unit that arranges the elements in the XML document acquired in a row according to the tree structure in an order of appearance in the XML document of the elements;
 - an extracting unit that extracts character-string data from each
 - 10 of the elements arranged, wherein the character-string data include character string in a start tag and an end tag, and character string within the start tag and the end tag in the elements;
 - an identifying unit that identifies a node-type of each of the elements arranged, wherein the node-type indicates a kind in the tree
 - 15 structure for each of the elements;
 - a generating unit that generates link data that associates character-string data extracted with a node-type identified for each of the elements arranged.
- 20 2. An XML documents retrieving device that retrieves character strings from character-string data in XML documents that have a predetermined tree structure, comprising:
 - an XML documents acquiring unit that acquires a plurality of
 - retrieval conditions;
 - 25 a documents-structure data judging unit that judges whether a

plurality of documents-structure data arranged according to the tree structure in an order of appearance of elements in the XML documents matches with a corresponding one of the respective retrieval conditions;

5 a related character-string extractor that extracts character strings in the character-string data associated with the documents-structure data that is judged to match with the retrieval conditions; and

a related character-string judging unit that judges whether character string extracted by the related character-string extractor
10 matches with a corresponding one of the retrieval conditions, wherein

when the related character-string judging unit judges that the character string matches with the retrieval conditions and when the documents-structure judging unit judges that the documents-structure data arranged subsequent to the documents-structure data that
15 matches with the retrieval conditions, the related character-string extracting unit extracts character-string in the character-string data associated with documents-structure data other than the documents-structure data that is judged to match with the retrieval conditions, as the character string subjected to retrieval.

20

3. A method of interpreting an XML document, comprising:
acquiring an XML document, wherein the XML document includes a plurality of elements that form a predetermined tree structure;

25 arranging the elements in the XML document acquired in a row

according to the tree structure in an order of appearance in the XML document of the elements;

extracting character-string data from each of the elements arranged, wherein the character-string data include character string in a start tag and an end tag, and character string within the start tag and the end tag in the elements;

identifying a node-type of each of the elements arranged, wherein the node-type indicates a kind in the tree structure for each of the elements;

generating link data that associates character-string data extracted with a node-type identified for each of the elements arranged.

4. The method according to claim 3, wherein the arranging includes placing child elements of a parent element between the parent element.

5. A method of retrieving an XML document in which a character string is retrieved from character-string data in the XML document that have a predetermined tree structure, comprising:

acquiring retrieval conditions;

judging whether first document-structure data obtained by arranging elements in the XML document according to the tree structure in an order of appearance of the elements matches with the retrieval conditions acquired;

extracting a character string in the character-string data

associated with the first document-structure data that is judged to match with the retrieval conditions;

judging whether the character string extracted matches with the retrieval conditions;

- 5 judging whether second document-structure data arranged subsequent to the first document-structure data matches with the retrieval conditions upon judging that the character string extracted matches with the retrieval conditions; and

- extracting a character string in the character-string data
10 associated with the second document-structure data that is judged to match with the retrieval conditions upon judging that the second document-structure data arranged subsequent to the first document-structure data matches with the retrieval conditions.

- 15 6. The method according to claim 5, wherein the first document-structure data and the second document-structure data include node-types that indicate kinds of nodes in the tree structure and link data that associates the character-string data.

- 20 7. The method according to claim 6, wherein
 the judging whether the first document-structure data matches with the retrieval conditions and the judging whether the second document-structure data matches with the retrieval conditions
 include judging whether the node-types and the link data match
25 with retrieval conditions.

8. The method according to claim 6, wherein
the extracting a character string in the character-string data
associated with the first document-structure data and the extracting a
5 character string in the character-string data associated with the second
document-structure data
include extracting the character-string data based on the link
data.
- 10 9. A computer program that realizes on a computer a method of
interpreting an XML document, the computer program making the
computer execute:
acquiring an XML document, wherein the XML document
includes a plurality of elements that form a predetermined tree
15 structure;
arranging the elements in the XML document acquired in a row
according to the tree structure in an order of appearance in the XML
document of the elements;
extracting character-string data from each of the elements
20 arranged, wherein the character-string data include character string in a
start tag and an end tag, and character string within the start tag and
the end tag in the elements;
identifying a node-type of each of the elements arranged,
wherein the node-type indicates a kind in the tree structure for each of
25 the elements;

generating link data that associates character-string data
extracted with a node-type identified for each of the elements arranged.

10. A computer program that realizes on a computer a method of
5 retrieving an XML document in which a character string is retrieved from
character-string data in the XML document that have a predetermined
tree structure, the computer program making the computer execute:
- acquiring retrieval conditions;
 - judging whether first document-structure data obtained by
10 arranging elements in the XML document according to the tree structure
in an order of appearance of the elements matches with the retrieval
conditions;
 - extracting a character string in the character-string data
associated with the first document-structure data that is judged to
15 match with the retrieval conditions;
 - judging whether the character string extracted matches with the
retrieval conditions;
 - judging whether second document-structure data arranged
subsequent to the first document-structure data matches with the
20 retrieval conditions upon judging that the character string extracted
matches with the retrieval conditions; and
 - extracting a character string in the character-string data
associated with the second document-structure data that is judged to
match with the retrieval conditions upon judging that the second
25 document-structure data arranged subsequent to the first

document-structure data matches with the retrieval conditions.